IN THE CLAIMS:

1. (Currently Amended) A video display device comprising:

a display configured to display a primary image and a picture-in-picture image (PIP) overlaying the primary image; and

a processor operatively coupled to the display and configured to receive a first video data stream for the primary image, to receive a second video data stream for the PIP, and to change a PIP display characteristic in response to a received audio ~~indication~~ command and a related gesture from a user.

2. (Original) The video display device of Claim 1, wherein the PIP display characteristic is at least one of a position of the PIP on the display and a display size of the PIP.

3. (Original) The video display device of Claim 1, comprising:

a microphone for receiving the audio indication from the user; and

a camera for acquiring an image of the user containing the related gesture.

4. (Original) The video display device of Claim 1, wherein the processor is configured to analyze audio information received from the user to identify when a PIP related audio indication is intended by the user.

5. (Original) The video display device of Claim 1, wherein the processor is configured to analyze image information received from the user after the audio indication is received to identify the change in the PIP display characteristic that is expressed by the received gesture.

-2-

6. (Original) The video display device of Claim 5, wherein the image information is contained in a sequence of images and wherein the processor is configured to analyze the sequence of images to determine the received gesture.

7. (Original) The video display device of Claim 1, wherein the image information is contained in a sequence of images and wherein the processor is configured to determine the received gesture by analyzing the sequence of images and determining a trajectory of a hand of the user.

8. (Original) The video display device of Claim 1, wherein the processor is configured to determine the received gesture by analyzing an image of the user and determining a posture of a hand of the user.

9. (Original) The video display device of Claim 1, wherein the video display device is a television.

10. (Original) The video display device of Claim 1, wherein the image is a sequence of images of the user containing the user gesture, the video display device comprising a camera for acquiring the sequence of images of the user.

11. (Original) A method of controlling a display characteristic of a picture-in-picture display (PIP) overlaying a primary display, the method comprising:

receiving an audio indication from a user;

determining whether the received audio indication is one of a plurality of expected audio indications;

analyzing a gesture of the user if the received audio indication is one of the plurality of expected audio indications; and

-3-

controlling the display characteristic if the gesture is a gesture related to the received audio indication.

12. (Original) The method of Claim 11, wherein analyzing the gesture comprises:

receiving a sequence of images; and

analyzing the sequence of images to determine the gesture.

13. (Original) The method of Claim 11, wherein analyzing the gesture comprises:

receiving a sequence of images;

analyzing the sequence of images to determine a trajectory of a hand of the user; and

determining the gesture by the determined trajectory.

14. (Original) The method of Claim 11, wherein analyzing the gesture comprises:

analyzing an image of the user to determine a posture of a hand of the user; and

determining the gesture by the determined posture .

15. (Original) A program segment stored on a processor readable medium for controlling a display characteristic of a picture-in-picture display (PIP) overlaying a primary display, the program segment comprising:

a program segment for controlling receipt of an audio indication;

a program segment for determining whether a received audio indication is one of a plurality of stored audio indications;

a program segment for analyzing a gesture of the user if the received audio indication is one of the plurality of stored audio indications; and

a program segment for controlling the display characteristic if the gesture is a gesture related to the received audio indication.

-4-

16. (Original) The program segment of Claim 15, wherein the program segment for analyzing the gesture comprises:

a program segment for controlling receipt of a sequence of images; and

a program segment for analyzing the sequence of images to determine the gesture.

17. (Original) The program segment of Claim 15, wherein the program segment for analyzing the gesture comprises:

a program segment for controlling receipt of a sequence of images;

a program segment for analyzing the sequence of images to determine a trajectory of a hand of the user; and

a program segment for determining the gesture by the determined trajectory.

18. (Original) The program segment of Claim 15, wherein the program segment for analyzing the gesture comprises:

a program segment for analyzing an image of the user to determine a posture of a hand of the user; and

a program segment for determining the gesture by the determined posture.

19. (New) A video display device comprising:

a display configured to display a primary image and a picture-in-picture image (PIP); and

a processor operatively coupled to the display and configured to receive a first video data stream for the primary image, to receive a second video data stream for the PIP, and to change a PIP display characteristic in response to a received audio indication and a related gesture from a user,

wherein the PIP display characteristic is at least one of a position of the PIP on the display and a

display size of the PIP.

20. (New) A video display device comprising:

a display configured to display a primary image and a picture-in-picture image (PIP);

and

a processor operatively coupled to the display and configured to receive a first video

data stream for the primary image, to receive a second video data stream for the PIP, and to change a

PIP display characteristic in response to a received audio indication and a related gesture from a user,

wherein the processor is configured to analyze image information received from the user after the

audio indication is received to identify the change in the PIP display characteristic that is expressed by

the received gesture.